

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 09/598,982C
Source: 1FW16
Date Processed by STIC: 8/10/05

ENTERED

Ed. Authorized by: J.E. C. Low

**CRF Errors Edited by the STIC Systems
Branch**

Serial Number: 09/598,982C

CRF Edit Date: 8/10/05
Edited by: SL

___ Realigned nucleic acid/amino acid numbers/text in cases where the sequence text "wrapped" to the next line

___ Corrected the SEQ ID NO. Sequence numbers edited were:

___ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited:

___ Deleted: ___ invalid beginning/end-of-file text ; ___ page numbers

___ Inserted mandatory headings/numeric identifiers, specifically:

___ Moved responses to same line as heading/numeric identifier, specifically:

✓ Other:

converted submitted file to a generated
Patent In sequence listing

Revised 09/09/2003



IFW16

RAW SEQUENCE LISTING

DATE: 08/10/2005

PATENT APPLICATION: US/09/598,982C

TIME: 10:50:00

Input Set : A:\34506104.ST25.txt

Output Set: N:\CRF4\08102005\I598982C.raw

```

3 <110> APPLICANT: Niles, Andrew
4     Maffitt, Mark
5     Haak-Frendscho, Mary
7 <120> TITLE OF INVENTION: RECOMBINANT PROTEOLYTIC TRYPTASES, ACTIVE SITE MUTANTS
THEREOF,
8     AND METHODS OF MAKING SAME
10 <130> FILE REFERENCE: 34506.104
12 <140> CURRENT APPLICATION NUMBER: 09/598,982C
13 <141> CURRENT FILING DATE: 2000-06-21
15 <150> PRIOR APPLICATION NUMBER: 09/079,970
16 <151> PRIOR FILING DATE: 1998-04-15
18 <160> NUMBER OF SEQ ID NOS: 52
20 <170> SOFTWARE: PatentIn version 3.3
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 735
24 <212> TYPE: DNA
25 <213> ORGANISM: Homo sapiens
28 <220> FEATURE:
29 <221> NAME/KEY: CDS
30 <222> LOCATION: (1)..(735)
32 <400> SEQUENCE: 1
33 atc gtc ggg ggt cag gag gcc ccc agg agc aag tgg ccc tgg cag gtg      48
34 Ile Val Gly Gly Gln Glu Ala Pro Arg Ser Lys Trp Pro Trp Gln Val
35 1          5          10          15
37 agc ctg aga gtc cac ggc cca tac tgg atg cac ttc tgc ggg ggc tcc      96
38 Ser Leu Arg Val His Gly Pro Tyr Trp Met His Phe Cys Gly Gly Ser
39          20          25          30
41 ctc atc cac ccc cag tgg gtg ctg acc gca gcg cac tgc gtg gga ccg      144
42 Leu Ile His Pro Gln Trp Val Leu Thr Ala Ala His Cys Val Gly Pro
43          35          40          45
45 gac gtc aag gat ctg gcc gcc ctc agg gtg caa ctg cgg gag cag cac      192
46 Asp Val Lys Asp Leu Ala Ala Leu Arg Val Gln Leu Arg Glu Gln His
47          50          55          60
49 ctc tac tac cag gac cag ctg ctg ccg gtc agc agg atc atc gtg cac      240
50 Leu Tyr Tyr Gln Asp Gln Leu Leu Pro Val Ser Arg Ile Ile Val His
51 65          70          75          80
53 cca cag ttc tac acc gcc cag atc gga gcg gac atc gcc ctg ctg gag      288
54 Pro Gln Phe Tyr Thr Ala Gln Ile Gly Ala Asp Ile Ala Leu Leu Glu
55          85          90          95
57 ctg gag gag ccg gtg aac gtc tcc agc cac gtc cac acg gtc acc ctg      336
58 Leu Glu Glu Pro Val Asn Val Ser Ser His Val His Thr Val Thr Leu
59          100          105          110
61 ccc cct gcc tca gag acc ttc ccc ccg ggg atg ccg tgc tgg gtc act      384
62 Pro Pro Ala Ser Glu Thr Phe Pro Pro Gly Met Pro Cys Trp Val Thr

```

RAW SEQUENCE LISTING

DATE: 08/10/2005

PATENT APPLICATION: US/09/598,982C

TIME: 10:50:00

Input Set : A:\34506104.ST25.txt

Output Set: N:\CRF4\08102005\I598982C.raw

```

63          115          120          125
65 ggc tgg ggc gat gtg gac aat gat gag cgc ctc cca ccg cca ttt cct      432
66 Gly Trp Gly Asp Val Asp Asn Asp Glu Arg Leu Pro Pro Pro Phe Pro
67          130          135          140
69 ctg aag cag gtg aag gtc ccc ata atg gaa aac cac att tgt gac gca      480
70 Leu Lys Gln Val Lys Val Pro Ile Met Glu Asn His Ile Cys Asp Ala
71 145          150          155          160
73 aaa tac cac ctt ggc gcc tac acg gga gac gac gtc cgc atc gtc cgt      528
74 Lys Tyr His Leu Gly Ala Tyr Thr Gly Asp Asp Val Arg Ile Val Arg
75          165          170          175
77 gac gac atg ctg tgt gcc ggg aac acc cgg agg gac tca tgc cag ggc      576
78 Asp Asp Met Leu Cys Ala Gly Asn Thr Arg Arg Asp Ser Cys Gln Gly
79          180          185          190
81 gac tcc gga ggg ccc ctg gtg tgc aag gtg aat ggc acc tgg ctg cag      624
82 Asp Ser Gly Gly Pro Leu Val Cys Lys Val Asn Gly Thr Trp Leu Gln
83          195          200          205
85 gcg ggc gtg gtc agc tgg ggc gag ggc tgt gcc cag ccc aac cgg cct      672
86 Ala Gly Val Val Ser Trp Gly Glu Gly Cys Ala Gln Pro Asn Arg Pro
87          210          215          220
89 ggc atc tac acc cgt gtc acc tac tac ttg gac tgg atc cac cac tat      720
90 Gly Ile Tyr Thr Arg Val Thr Tyr Tyr Leu Asp Trp Ile His His Tyr
91 225          230          235          240
93 gtc ccc aaa aag ccg      735
94 Val Pro Lys Lys Pro
95          245
98 <210> SEQ ID NO: 2
99 <211> LENGTH: 245
100 <212> TYPE: PRT
101 <213> ORGANISM: Homo sapiens
103 <400> SEQUENCE: 2
105 Ile Val Gly Gly Gln Glu Ala Pro Arg Ser Lys Trp Pro Trp Gln Val
106 1          5          10          15
109 Ser Leu Arg Val His Gly Pro Tyr Trp Met His Phe Cys Gly Gly Ser
110          20          25          30
113 Leu Ile His Pro Gln Trp Val Leu Thr Ala Ala His Cys Val Gly Pro
114          35          40          45
117 Asp Val Lys Asp Leu Ala Ala Leu Arg Val Gln Leu Arg Glu Gln His
118          50          55          60
121 Leu Tyr Tyr Gln Asp Gln Leu Leu Pro Val Ser Arg Ile Ile Val His
122 65          70          75          80
125 Pro Gln Phe Tyr Thr Ala Gln Ile Gly Ala Asp Ile Ala Leu Leu Glu
126          85          90          95
129 Leu Glu Glu Pro Val Asn Val Ser Ser His Val His Thr Val Thr Leu
130          100          105          110
133 Pro Pro Ala Ser Glu Thr Phe Pro Pro Gly Met Pro Cys Trp Val Thr
134          115          120          125
137 Gly Trp Gly Asp Val Asp Asn Asp Glu Arg Leu Pro Pro Pro Phe Pro
138          130          135          140
141 Leu Lys Gln Val Lys Val Pro Ile Met Glu Asn His Ile Cys Asp Ala

```

RAW SEQUENCE LISTING

DATE: 08/10/2005

PATENT APPLICATION: US/09/598,982C

TIME: 10:50:00

Input Set : A:\34506104.ST25.txt

Output Set: N:\CRF4\08102005\I598982C.raw

```

142 145          150          155          160
145 Lys Tyr His Leu Gly Ala Tyr Thr Gly Asp Asp Val Arg Ile Val Arg
146          165          170          175
149 Asp Asp Met Leu Cys Ala Gly Asn Thr Arg Arg Asp Ser Cys Gln Gly
150          180          185          190
153 Asp Ser Gly Gly Pro Leu Val Cys Lys Val Asn Gly Thr Trp Leu Gln
154          195          200          205
157 Ala Gly Val Val Ser Trp Gly Glu Gly Cys Ala Gln Pro Asn Arg Pro
158          210          215          220
161 Gly Ile Tyr Thr Arg Val Thr Tyr Tyr Leu Asp Trp Ile His His Tyr
162 225          230          235          240
165 Val Pro Lys Lys Pro
166          245
169 <210> SEQ ID NO: 3
170 <211> LENGTH: 40
171 <212> TYPE: DNA
172 <213> ORGANISM: Artificial Sequence
174 <220> FEATURE:
175 <223> OTHER INFORMATION: PCR primer
177 <400> SEQUENCE: 3
178 gggccctcg agaaaagaat cgtcgggggt caggaggccc          40
181 <210> SEQ ID NO: 4
182 <211> LENGTH: 40
183 <212> TYPE: DNA
184 <213> ORGANISM: Artificial Sequence
186 <220> FEATURE:
187 <223> OTHER INFORMATION: PCR primer
189 <400> SEQUENCE: 4
190 ccactatgtc ccaaaaaagc cgtgaagcgg ccgccgtcgt          40
193 <210> SEQ ID NO: 5
194 <211> LENGTH: 771
195 <212> TYPE: DNA
196 <213> ORGANISM: Homo sapiens
199 <220> FEATURE:
200 <221> NAME/KEY: CDS
201 <222> LOCATION: (7)..(753)
203 <400> SEQUENCE: 5
204 gggccc ctc gag aaa aga atc gtc ggg ggt cag gag gcc ccc agg agc          48
205      Leu Glu Lys Arg Ile Val Gly Gly Gln Glu Ala Pro Arg Ser
206      1          5          10
208 aag tgg ccc tgg cag gtg agc ctg aga gtc cac ggc cca tac tgg atg          96
209 Lys Trp Pro Trp Gln Val Ser Leu Arg Val His Gly Pro Tyr Trp Met
210 15          20          25          30
212 cac ttc tgc ggg ggc tcc ctc atc cac ccc cag tgg gtg ctg acc gca          144
213 His Phe Cys Gly Gly Ser Leu Ile His Pro Gln Trp Val Leu Thr Ala
214          35          40          45
216 gcg cac tgc gtg gga ccg gac gtc aag gat ctg gcc gcc ctc agg gtg          192
217 Ala His Cys Val Gly Pro Asp Val Lys Asp Leu Ala Ala Leu Arg Val
218          50          55          60

```

RAW SEQUENCE LISTING

DATE: 08/10/2005

PATENT APPLICATION: US/09/598,982C

TIME: 10:50:00

Input Set : A:\34506104.ST25.txt

Output Set: N:\CRF4\08102005\I598982C.raw

```

220  caa  ctg  cgg  gag  cag  cac  ctc  tac  tac  cag  gac  cag  ctg  ctg  ccg  gtc      240
221  Gln  Leu  Arg  Glu  Gln  His  Leu  Tyr  Tyr  Gln  Asp  Gln  Leu  Leu  Pro  Val
222           65                        70                        75
224  agc  agg  atc  atc  gtg  cac  cca  cag  ttc  tac  acc  gcc  cag  atc  gga  gcg      288
225  Ser  Arg  Ile  Ile  Val  His  Pro  Gln  Phe  Tyr  Thr  Ala  Gln  Ile  Gly  Ala
226      80                        85                        90
228  gac  atc  gcc  ctg  ctg  gag  ctg  gag  gag  ccg  gtg  aac  gtc  tcc  agc  cac      336
229  Asp  Ile  Ala  Leu  Leu  Glu  Leu  Glu  Glu  Pro  Val  Asn  Val  Ser  Ser  His
230  95                        100                        105                        110
232  gtc  cac  acg  gtc  acc  ctg  ccc  cct  gcc  tca  gag  acc  ttc  ccc  ccg  ggg      384
233  Val  His  Thr  Val  Thr  Leu  Pro  Pro  Ala  Ser  Glu  Thr  Phe  Pro  Pro  Gly
234           115                        120                        125
236  atg  ccg  tgc  tgg  gtc  act  ggc  tgg  ggc  gat  gtg  gac  aat  gat  gag  cgc      432
237  Met  Pro  Cys  Trp  Val  Thr  Gly  Trp  Gly  Asp  Val  Asp  Asn  Asp  Glu  Arg
238           130                        135                        140
240  ctc  cca  ccg  cca  ttt  cct  ctg  aag  cag  gtg  aag  gtc  ccc  ata  atg  gaa      480
241  Leu  Pro  Pro  Pro  Phe  Pro  Leu  Lys  Gln  Val  Lys  Val  Pro  Ile  Met  Glu
242      145                        150                        155
244  aac  cac  att  tgt  gac  gca  aaa  tac  cac  ctt  ggc  gcc  tac  acg  gga  gac      528
245  Asn  His  Ile  Cys  Asp  Ala  Lys  Tyr  His  Leu  Gly  Ala  Tyr  Thr  Gly  Asp
246      160                        165                        170
248  gac  gtc  cgc  atc  gtc  cgt  gac  gac  atg  ctg  tgt  gcc  ggg  aac  acc  cgg      576
249  Asp  Val  Arg  Ile  Val  Arg  Asp  Asp  Met  Leu  Cys  Ala  Gly  Asn  Thr  Arg
250  175                        180                        185                        190
252  agg  gac  tca  tgc  cag  ggc  gac  tcc  gga  ggg  ccc  ctg  gtg  tgc  aag  gtg      624
253  Arg  Asp  Ser  Cys  Gln  Gly  Asp  Ser  Gly  Gly  Pro  Leu  Val  Cys  Lys  Val
254           195                        200                        205
256  aat  ggc  acc  tgg  ctg  cag  gcg  ggc  gtg  gtc  agc  tgg  ggc  gag  ggc  tgt      672
257  Asn  Gly  Thr  Trp  Leu  Gln  Ala  Gly  Val  Val  Ser  Trp  Gly  Glu  Gly  Cys
258           210                        215                        220
260  gcc  cag  ccc  aac  cgg  cct  ggc  atc  tac  acc  cgt  gtc  acc  tac  tac  ttg      720
261  Ala  Gln  Pro  Asn  Arg  Pro  Gly  Ile  Tyr  Thr  Arg  Val  Thr  Tyr  Tyr  Leu
262      225                        230                        235
264  gac  tgg  atc  cac  cac  tat  gtc  ccc  aaa  aag  ccg  tgaagcggcc gccgtcgt      771
265  Asp  Trp  Ile  His  His  Tyr  Val  Pro  Lys  Lys  Pro
266      240                        245
269 <210> SEQ ID NO: 6
270 <211> LENGTH: 249
271 <212> TYPE: PRT
272 <213> ORGANISM: Homo sapiens
274 <400> SEQUENCE: 6
276  Leu  Glu  Lys  Arg  Ile  Val  Gly  Gly  Gln  Glu  Ala  Pro  Arg  Ser  Lys  Trp
277  1           5           10           15
280  Pro  Trp  Gln  Val  Ser  Leu  Arg  Val  His  Gly  Pro  Tyr  Trp  Met  His  Phe
281           20           25           30
284  Cys  Gly  Gly  Ser  Leu  Ile  His  Pro  Gln  Trp  Val  Leu  Thr  Ala  Ala  His
285      35           40           45
288  Cys  Val  Gly  Pro  Asp  Val  Lys  Asp  Leu  Ala  Ala  Leu  Arg  Val  Gln  Leu
289      50           55           60

```

RAW SEQUENCE LISTING

DATE: 08/10/2005

PATENT APPLICATION: US/09/598,982C

TIME: 10:50:00

Input Set : A:\34506104.ST25.txt

Output Set: N:\CRF4\08102005\I598982C.raw

```

292 Arg Glu Gln His Leu Tyr Tyr Gln Asp Gln Leu Leu Pro Val Ser Arg
293 65              70              75              80
296 Ile Ile Val His Pro Gln Phe Tyr Thr Ala Gln Ile Gly Ala Asp Ile
297              85              90              95
300 Ala Leu Leu Glu Leu Glu Glu Pro Val Asn Val Ser Ser His Val His
301              100              105              110
304 Thr Val Thr Leu Pro Pro Ala Ser Glu Thr Phe Pro Pro Gly Met Pro
305              115              120              125
308 Cys Trp Val Thr Gly Trp Gly Asp Val Asp Asn Asp Glu Arg Leu Pro
309              130              135              140
312 Pro Pro Phe Pro Leu Lys Gln Val Lys Val Pro Ile Met Glu Asn His
313 145              150              155              160
316 Ile Cys Asp Ala Lys Tyr His Leu Gly Ala Tyr Thr Gly Asp Asp Val
317              165              170              175
320 Arg Ile Val Arg Asp Asp Met Leu Cys Ala Gly Asn Thr Arg Arg Asp
321              180              185              190
324 Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys Lys Val Asn Gly
325              195              200              205
328 Thr Trp Leu Gln Ala Gly Val Val Ser Trp Gly Glu Gly Cys Ala Gln
329              210              215              220
332 Pro Asn Arg Pro Gly Ile Tyr Thr Arg Val Thr Tyr Tyr Leu Asp Trp
333 225              230              235              240
336 Ile His His Tyr Val Pro Lys Lys Pro
337              245
340 <210> SEQ ID NO: 7
341 <211> LENGTH: 27
342 <212> TYPE: DNA
343 <213> ORGANISM: Artificial Sequence
345 <220> FEATURE:
346 <223> OTHER INFORMATION: Mutagenesis oligonucleotide
348 <400> SEQUENCE: 7
349 gaggagccgg tgaaggtctc cagccac
352 <210> SEQ ID NO: 8
353 <211> LENGTH: 771
354 <212> TYPE: DNA
355 <213> ORGANISM: Homo sapiens
358 <220> FEATURE:
359 <221> NAME/KEY: CDS
360 <222> LOCATION: (7)..(753)
362 <400> SEQUENCE: 8
363 gggccc ctc gag aaa aga atc gtc ggg ggt cag gag gcc ccc agg agc
364      Leu Glu Lys Arg Ile Val Gly Gly Gln Glu Ala Pro Arg Ser
365      1              5              10
367 aag tgg ccc tgg cag gtg agc ctg aga gtc cac ggc cca tac tgg atg
368 Lys Trp Pro Trp Gln Val Ser Leu Arg Val His Gly Pro Tyr Trp Met
369 15              20              25              30
371 cac ttc tgc ggg ggc tcc ctc atc cac ccc cag tgg gtg ctg acc gca
372 His Phe Cys Gly Gly Ser Leu Ile His Pro Gln Trp Val Leu Thr Ala
373              35              40              45

```

VERIFICATION SUMMARY

DATE: 08/10/2005

PATENT APPLICATION: US/09/598,982C

TIME: 10:50:01

Input Set : A:\34506104.ST25.txt

Output Set: N:\CRF4\08102005\I598982C.raw

**Raw Sequence Listing before editing,
for reference only**

09/598,982C

yy0000CProject0000yy00

(sample of submitted file)

Does Not Comply
Corrected Diskette Needed